



東北大学における内製生成AI基盤 ～生成AI革新期における実践報告～

国立大学法人東北大学 情報部デジタル変革推進課
デジタルイノベーションユニット 専門職員

大学院情報科学研究科博士後期課程 菅沼研究室

鈴木 翔太



DX推進のフェーズ

2020

業務のDX推進プロジェクト始動

- AIチャットボット稼働
- テレワークの制度化
- 電子決裁の開始

2021

コロナ対応と共に進むDX

- リーガルテック導入
- メタバース表彰式開催

2022

先端技術の積極的な業務応用

- AIナレーター導入
- メタバースにてDX展示会場を作成

2023

生成AIを大学業務に活用

- 全国の大学に先駆けて生成AIを導入

2024

生成AIのさらなる応用

- 生成AI対応チャットボットを
国立大学で初めて導入
- 継続的な生成AI研修の開催
- Tohoku University GAI

コロナ禍を契機にDXを加速し、
先端技術の積極的な活用で業務改革を継続的にアップデート





生成AIと東北大学

2022年	11月30日	ChatGPT公開
	12月4日	ユーザー数100万突破 ^①
2023年	1月	ユーザー数1億突破 ^②
	2月1日	ChatGPT Plus発表
	3月15日	GPT-4公開
	3月31日	<u>「ChatGPT等の生成系AI利用に関する留意事項」通知^③</u>
	5月18日	<u>プレスリリース「全国の大学に先駆けて生成AIサービスを導入 ～AIを駆使しDXの最先端を切り拓く」^④</u>
2024年	3月29日	<u>プレスリリース「生成 AI 対応チャットボットを 国立大学法人で初めて導入」^⑤</u>
	11月1日	<u>「学内専用生成AIサービスTohoku University GAI」の運用開始</u>



①



②



③



④



⑤



全国の大学に先駆けて生成AIサービスを導入

- ・対象業務や組織を限定せずにアカウントを配布し、生成AIをまずは使ってみる
- ・利用者アンケートにより活用容態および必要機能等を検討

本学での生成 AI サービス利用者アンケート結果
(生成 AI 活用可能業務)

生成 AI 活用可能業務 (シングルモーダル)
文章校正 (通知文など)
文章作成補助 (通知文など)
要約 (議事録など)
論点整理
業務報告書作成 (日報・分報) から所定様式への変換
コーディング・コードレビュー
作業手順調査
データ分析の補助
特許明細書作成
英会話学習
英訳補助
専門用語理解の補助



指示および入力データ形式による生成 AI 活用類型

類型	指示	元となる データ形式	例
1	作成・加工	テキスト	文章校正など
2	作成・加工	非テキスト	会議事録作成など
3	教示	テキスト	専門用語理解の補助など
4	教示	非テキスト	英会話学習など

各類型に応じて、更なる業務活用方策を検討

導入した生成AIサービス（テキストシングルモーダル）で効率化実現達成

非テキストデータをテキストデータに変換する機構を生成AIに接続することで、更なる業務効率化が可能 ⇒ **マルチモーダル化**

LLMの知識を拡張することで、更なる業務効率化が可能 ⇒ **RAG**



マルチモーダル議事録作成アプリの内製

対象業務の選定

部署横断的に発生する業務であり、かつ発生頻度も高い会議事録作成業務を対象として検証

要約作成アプリ利用者アンケート集計結果

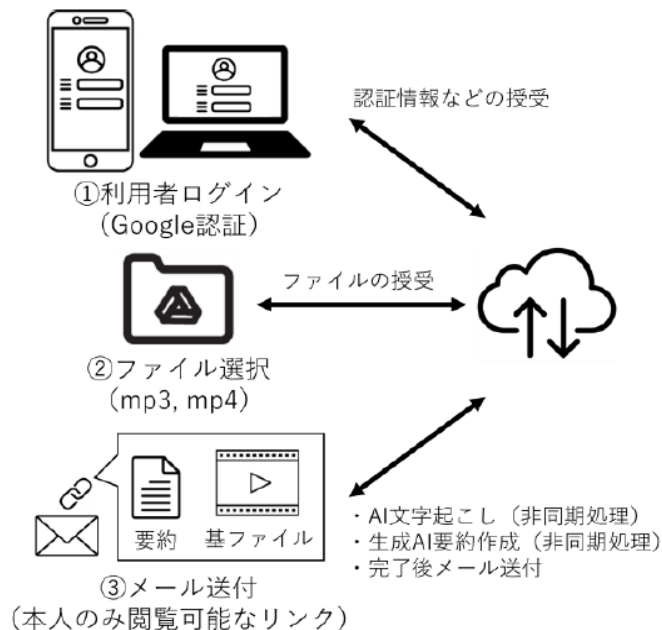
項目	数値
アプリの月間平均利用回数	3.63 回 / 月・人
アプリ未利用時の平均要約作成所要時間 (オンライン会議 1 時間あたり)	1.98 時間 / 人
アプリ利用時の平均要約作成所要時間 (オンライン会議 1 時間あたり)	0.51 時間 / 人

本アプリの特徴

- 音声ファイル (mp3) のみならず動画ファイル (mp4) から作成可能
- 入力ファイル容量制限なし
- 自組織システム基盤内 (Google) に閉じたアプリのため、高セキュリティ
- 内製のため費用が安価 (年間約10万円)

要約作成アプリ開発環境および活用 AI モデル

機能	環境・モデル
プラットフォーム	Google App Engine
開発言語	Python
音声言語モデル	Whisper
生成 AI モデル	GPT-4 Turbo



要約作成システムの利用フロー



RAGの活用（チャットボット）

6



東北大学には
どんなことが
学べるの？

こんなことが
学べるハギ

ここの処理は
こうすると
いいハギ

この業務はど
う処理する？



全学チャットボット

学内**約350のWebページ**から検索して回答

受験生や
地域の方など



業務マニュアルチャットボット

多様な形式の業務マニュアルを一括検索して回答

職員



手続きが分
からないだけ
ど・・・

ここにアク
セスして手続
きするハギ

こんな研究
をしているハギ

工学研究科
の情報を詳し
く知りたい

在学生向けチャットボット

学生と共同で運用し、**LINE**で情報提供



在学生



部局チャットボット

利用希望部署にもシステム環境を配布し
独自のチャットボットを構築



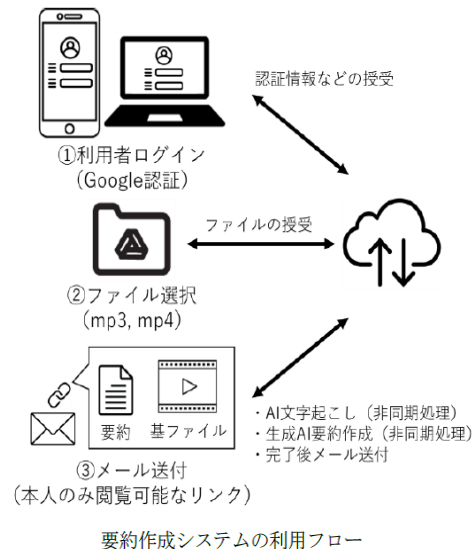
入学希望者



シングルモードルによる 生成AI活用の実験的検証

表 1 本学での生成 AI サービス利用者アンケート結果
(生成 AI 活用可能業務)

生成 AI 活用可能業務 (シングルモードル)
文章校正 (通知文など)
文章作成補助 (通知文など)
要約 (議事録など)
論点整理
業務報告書作成 (日報・分報) から所定様式への変換
コーディング・コードレビュー
作業手順調査
データ分析の補助
特許明細書作成
英会話学習
英訳補助
専門用語理解の補助



マルチモードル化による要約作成効率化

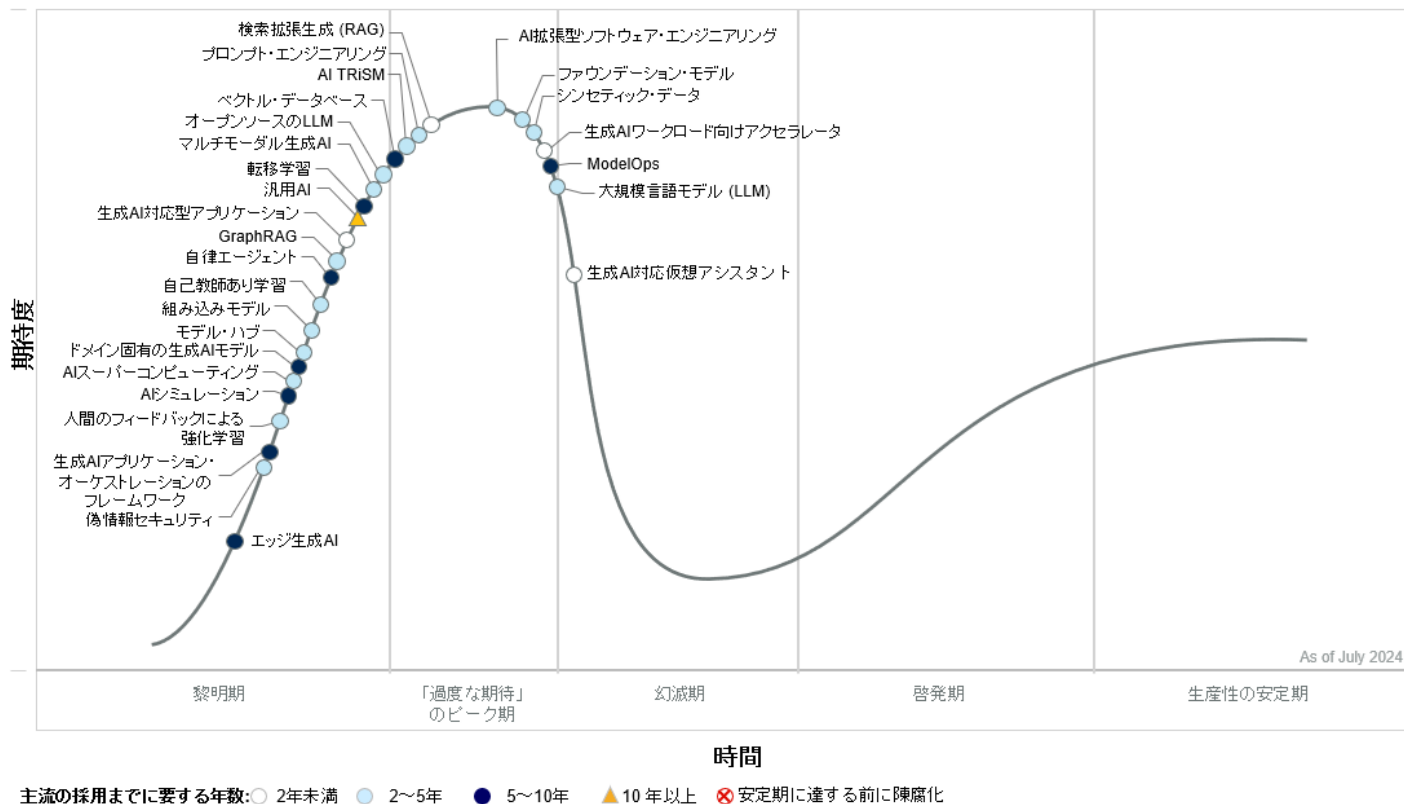


様々な業務へRAGを活用



ダイナミックに変動する生成AI関連技術

8



ガートナー「生成AIのハイプ・サイクル：2024

<https://www.gartner.co.jp/ja/newsroom/press-releases/pr-20240910-genai-hc>

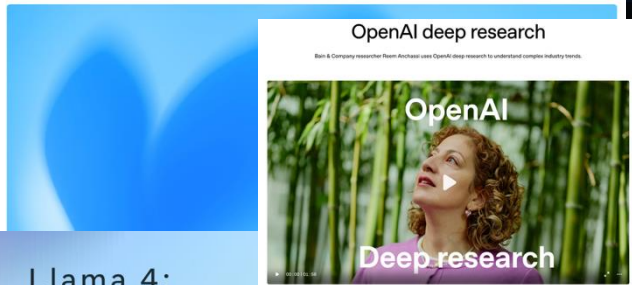


ダイナミックに変動する生成AI業界

9



Evolving OpenAI's structure



Llama 4: Leading Multimodal Intelligence

Newest model suite offering unprecedented speed and efficiency

Llama 4 Behemoth

288B active parameter, 16 experts
2T total parameters

The most intelligent teacher model for distillation

Preview

Llama 4 Maverick

17B active parameters, 128 experts
400B total parameters

Native multimodal with 1M context length

Available

Llama 4 Scout

17B active parameters, 16 experts
109B total parameters

Industry leading 10M context length
Optimized inference

Available

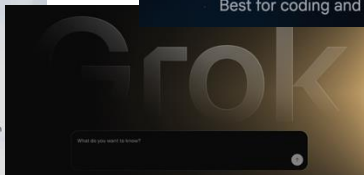


	Claude Opus 4	Claude Sonnet 4	Claude Sonnet 3.7	OpenAI o3	OpenAI GPT-4.1	Gemini 2.5 Pro (Preview) (05-06)
Agentic coding SWE-bench Verified ^{1,2}	72.5% / 79.4%	72.7% / 80.2%	62.3% / 70.3%	69.1%	54.6%	63.2%
Agentic terminal coding Terminal-bench ^{3,4}	43.2% / 50.0%	35.5% / 41.3%	35.2%	30.2%	30.3%	25.3%
Graduate-level reasoning GPQA Diamond ⁵	79.6% / 83.3%	75.4% / 83.8%	78.2%	83.3%	66.3%	83.0%
Agentic tool use TASU-bench	Retail 81.4% Airline 59.6%	Retail 80.5% Airline 60.0%	Retail 81.2% Airline 58.4%	Retail 70.4% Airline 52.0%	Retail 68.0% Airline 49.4%	—
Multilingual Q&A MMMU (validation)	88.8%	86.5%	85.9%	88.8%	83.7%	—
Visual reasoning MMMU (validation)	76.5%	74.4%	75.0%	82.9%	74.8%	79.6%
High school math competition AIME 2025 ^{6,7}	75.5% / 90.0%	70.5% / 85.0%	54.8%	88.9%	—	—

Methodology
1. OpenAI and Sonnet 4 achieved 72.5% and 72.7% respectively on SWE-bench Verified across 10 trials, single attempt problems, no test-time compute, using Claude Sonnet 3.7 as the baseline.
2. OpenAI and Sonnet 4 achieved 43.2% and 35.5% respectively on Terminal-bench across 10 trials, single attempt problems, no test-time compute, using Claude Sonnet 3.7 as the baseline.
3. OpenAI and Sonnet 4 achieved 79.6% and 75.4% respectively on GPQA Diamond across 10 trials, single attempt problems, no test-time compute, using Claude Sonnet 3.7 as the baseline.
4. OpenAI and Sonnet 4 achieved 81.4% and 80.5% respectively on TASU-bench across 10 trials, single attempt problems, no test-time compute, using Claude Sonnet 3.7 as the baseline.
5. OpenAI and Sonnet 4 achieved 88.8% and 86.5% respectively on MMU (validation) across 10 trials, single attempt problems, no test-time compute, using Claude Sonnet 3.7 as the baseline.
6. OpenAI and Sonnet 4 achieved 75.5% and 70.5% respectively on AIME 2025 across 10 trials, single attempt problems, no test-time compute, using Claude Sonnet 3.7 as the baseline.
7. OpenAI and Sonnet 4 achieved 75.5% and 70.5% respectively on AIME 2025 across 10 trials, single attempt problems, no test-time compute, using Claude Sonnet 3.7 as the baseline.

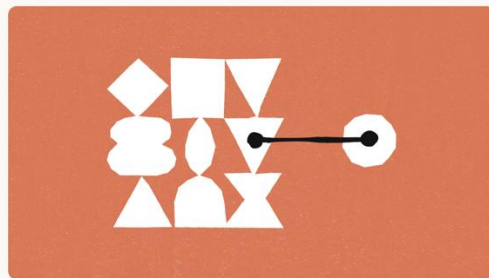
Gemini 2.5 Pro

Best for coding and complex prompts



Introducing the Model Context Protocol

2024年11月26日 · 3 min read



deepseek

Qwen3-235B-A228 MoE	Qwen3-32B Dense	OpenAI-o1 2024-12-17	Deepseek-R1	Grok 3 Beta Think	Gemini2.5-Pro	OpenAI-o3-mini Medium
95.6	93.8	92.1	93.2	-	96.4	89.0
85.7	81.4	74.3	79.8	83.9	92.0	79.6
81.5	72.9	79.2	70.0	77.3	86.7	74.8
70.7	65.7	63.9	64.3	70.6	70.4	66.3
2056	1977	1891	2029	-	2001	2036
61.8	50.2	61.7	56.9	53.3	72.9	53.8
77.1	74.9	75.7	71.6	-	82.4	70.0
70.8	70.3	67.8	56.9	-	62.9	64.6
71.9	73.0	48.8	67.7	-	77.8	48.4

1. AIME 24-25: We sample 40 times for each group and report the average of the accuracy. AIME 24 consists of Part I and Part II, with a total of 30 questions.
2. AIME: We didn't achieve the best results of Qwen3 in terms of efficiency and performance.
3. BFCL: The Qwen3 results are evaluated using the BFCL format, while the baseline results are assessed using the highest scores obtained from either the BFCL or prompt format.



- ・全教職員が安心して利用可能な生成AIシステムを構築
- ・AWS社のOSS(Generative AI Use Cases)をカスタマイズして開発

The screenshot shows the Tohoku University GAI web application. The sidebar on the left contains navigation links for various features. The main area displays a grid of use cases, each with an icon, a title, a brief description, and a 'Try' button. The use cases include Chat, Agent Chat, Document Generation, Summary, Correction, Translation, Web Content Extraction, and Image Generation.

利用実績

アカウント数	約1,000名
平均入力トークン数 (daily)	約2.5M token
音声認識時間(daily)	約20時間
利用料 (monthly)	\$2,500

システム構成

クラウド基盤	AWS
生成AIサービス	Amazon Bedrock
音声認識サービス	Amazon Transcribe
利用可能なLLM	Claude 4.0 Sonnetなど8種
検索エンジン (Agent用)	Google Custom Search API



ユースケース (生成 AI)

- ホーム
- チャット
- 文章生成
- 要約
- 執筆
- 翻訳
- Web コンテンツ抽出
- 画像生成
- 映像分析
- ダイアグラム生成

ツール (AI サービス)

- 音声認識

会話履歴

件名で検索

- Japanese to English Name Intr
- 大学における生成AI導入と活用
- Japanese to English Translatio
- Japanese to English Translatio
- Japanese to English Translatio
- Translation Request for Directi
- Japanese to English Translatio
- Japanese to English Translatio
- Japanese to English Translatio
- Japanese to English Translatio
- Japanese to English Translatio
- Japanese to English Translatio
- Japanese to English Translatio

shota.suzuki.e8@tohoku.ac.jp

チャット

Claude Sonnet 4



プロンプト例を見る

システムプロンプト

入力してください



ユースケース (生成 AI)

ホーム

チャット

文章生成

要約

執筆

翻訳

Web コンテンツ抽出

画像生成

映像分析

ダイアグラム生成

ツール (AI サービス)

音声認識

会話履歴

件名で検索

東北大学における生成AI活用事

東北大学についての関西での

東北大学の基本情報と特徴につ

Japanese to English Name Intr

大学における生成AI導入と活用

Japanese to English Translatio

Japanese to English Translatio

Japanese to English Translatio

Translation Request for Directi

Japanese to English Translatio

Japanese to English Translatio

Japanese to English Translatio

shota.suzuki.e8@tohoku.ac.jp

音声認識

マイク入力 or ファイルアップロードから選択してください

マイク入力

録音を開始する

ファイルアップロード

ファイルを選択

選択されていません

mp3, mp4, wav, flac, ogg, amr, webm, m4a ファイルが利用可能です

詳細なパラメータ

クリア

実行

音声認識結果がここに表示されます



ユースケース (生成 AI)

- ホーム
- チャット
- 文章生成
- 要約
- 執筆
- 翻訳**
- Web コンテンツ抽出
- 画像生成
- 映像分析
- ダイアグラム生成

ツール (AI サービス)

- 音声認識

会話履歴

件名で検索

- 東北大学情報部係長会議AI活用
- 東北大学における生成AI活用事
- 東北大学についての関西弁での
- 東北大学の基本情報と特徴につ
- Japanese to English Name Intr
- 大学における生成AI導入と活用
- Japanese to English Translatio
- Japanese to English Translatio
- Japanese to English Translatio
- Translation Request for Directi
- Japanese to English Translatio
- Japanese to English Translatio

shota.suzuki.e8@tohoku.ac.jp

翻訳

翻訳したい文章

Claude Sonnet 4

自動翻訳

言語を自動検出

英語

入力してください

翻訳結果がここに表示されます

クリア 実行

追加コンテキスト - 任意



ユースケース (生成 AI)

ホーム

チャット

文章生成

要約

執筆

翻訳

Web コンテンツ抽出

画像生成

映像分析

ダイアグラム生成

ツール (AI サービス)

音声認識

会話履歴

件名で検索

卵焼きの作り方フローチャート

経費清算システムのフローチャ

会社経費精算フローチャート作

CSVデータ処理システムのフロ

CSV処理システムのフロー図作

人物評価についての相談

Japanese to English Translatio

Japanese to English Translatio

東北大学情報部係長会議AI活用

東北大学における生成AI活用例

東北大学についての関西井での

東北大学の基本情報と経緯につ

shota.suzuki.e8@tohoku.ac.jp

ダイアグラム生成

ダイアグラム生成元の文章

モデル

Claude Sonnet 4

図の種類を選択 - AI

AI

AIに最適なダイアグラムタイプを選択させます

フローチャート

フローチャートはプロセスやワークフローを視覚化します

円グラフ

円グラフは全体に対する割合を表現します

マインドマップ

マインドマップは関連する概念やアイデアを階層的に表現します

他の図の種類を見る

生成元の文章を入力してください

入力例

経費フロー例

クリア

生成

生成結果

ダイアグラムがここに表示されます



ユースケース (生成 AI)

ホーム

チャット

SearchEngine Agent

CodeInterpreter Agent

文章生成

要約

執筆

翻訳

Web コンテンツ抽出

画像生成

映像分析

デザインAI生成

ツール (AI サービス)

音声認識

会話履歴

件名で検索

銀河系画像生成プロンプト作成

可愛い猫の画像生成プロンプト

Japanese to English Translatio

名前の誤解を否定する記事要約

東北大学の生成AI活用論文につ

東北大学の基本情報と特徴につ

東北大学の基本情報と特徴につ

DifyカスタムツールでAPIキーを

DifyでOpenAPIの固定パラメー

GPTsでWordPress投稿botを作

DifyツールのWordPress API機

記事生成・抽出・分析ツール

shota.suzuki.e8@tohoku.ac.jp

画像生成アシスタント

Claude Sonnet 4

ヒント

「かわいい猫の画像を生成して」のような簡単な指示から始めてみましょう。

「水彩画風」や「リアルに」などスタイルを指定すると効果的です。

「背景に山」や「青空」など、含めたい要素を詳細に描写しましょう。

「高解像度」や「細部まで描写」など、品質に関する指示も効果的です。

「アニメ風」や「油絵風」など、特定の芸術スタイルを指定することもできます。

複雑なシーンの場合は、構図、配色などについて具体的に指示しましょう。

どのような画像を生成したいですか？

プロンプト

入力してください

ネガティブプロンプト

入力してください

モデル

Stable Image Ultra v1.1

解像度

1024 x 1024

シード

0

画像サンプル数

3

詳細なパラメータ

生成

クリア

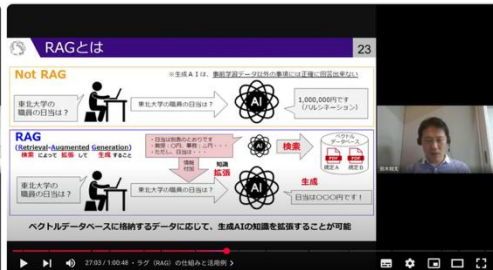


定期的な生成AIセミナー開催

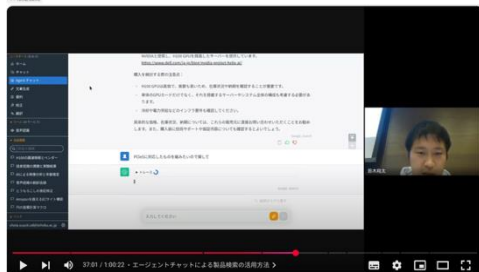
生成AIに関する情報の発信



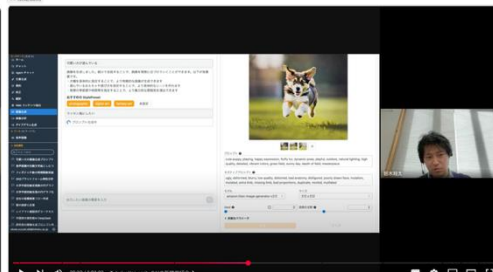
第一回生成AIセミナー（東北大学業務のDX推進プロジェクト・チーム）



第二回生成AIセミナー（東北大学業務のDX推進プロジェクト・チーム）



第三回生成AIセミナー（東北大学業務のDX推進プロジェクト・チーム）



第四回生成AIセミナー（東北大学業務のDX推進プロジェクト・チーム）

- ・これまでオンラインにて4回開催
- ・毎回約300人の教職員が参加



過去のセミナー資料・動画や
学内で利用可能なサービス等を掲載



BENEFIT : 有機的なソリューションプールの利活用

DX推進に効く事例の利活用

DXやそれに伴うBXの事例が多数存在します。
それらを利活用して各機関のDX推進にお役立てください。



共創による新たな可能性

教育研究機関に加え、企業も参画しています。
組織の枠を超えた共創によるあらたな可能性があります。

機関固有の悩み・解決策の共有



国公私立問わず多くの機関が参画しています。
組織固有の悩みや解決策を共有できます。



(2025年6月現在)

73 大学 1 行政機関 30 企業 2 団体 から 422 名が参加



CONTACT : お問い合わせ・参加希望



お問い合わせや参加希望についてはどうぞお気軽に、こちらからご連絡ください。
みなさまとお会いできることを楽しみにしております。